Degraded Service Event

Event Period:	3/16/10 4:45pm EST - 3/16/10 10:45pm EST
System(s) Affected:	Operations/Partner Test/Testbed
Product(s) Affected:	WIST/ECHO/Ingest/Website

Executive Summary:

All instances of ECHO and WIST experienced an outage on the evening of Tuesday 3/16/10 starting at 4:45pm EST. Analysis has identified that a significant network event occurred which caused a loss of internal and external network connectivity. ECHO Operations was notified of the system issues by its automated monitoring tool and immediately contacted ECHO System Administrators. ECHO System Administrators were able to restore connectivity to all systems by 6:00pm EST. At that point, the Partner Test and Testbed systems were fully available. The Operational system remained unavailable due to an issue with connectivity between the Oracle RAC nodes. This problem was identified at 7:00pm EST and the ECHO System Administration and Database teams worked to identify the root cause of the problem. By 9:00pm EST the decision was made to restart each RAC node and restore them to proper working order. This took approximately 1 hour, completing at 10:00pm EST. Subsequent to this activity, the Operational kernels were restarted, restoring ECHO and WIST search and order capabilities. Operational Ingest was restarted at 10:45pm EST. No data loss or corruption occurred as a result of this outage.

Detailed Summary:

On January 13th of this year, a redundant network link from the ECHO network to the GSFC EBNET network was installed. This introduced a redundant network path from ECHO in order to increase fault tolerance for network issues. On 1/19/10 at approximately 10:50am, ECHO became the target a Distributed Denial of Service (DDoS) which caused significant connectivity issues to the ECHO system. ECHO System Administrators worked very closely with the GSFC network security team to restore connectivity to the ECHO system. Connectivity was restored within approximately 1 day. Since that time brief network outages were experienced, but appeared to be related to the DDoS attack.

Staring on Friday, 3/14/10, ECHO automated monitoring began reporting connectivity losses from the Operational kernels to the ECHO Data Provider order fulfillment endpoints. These errors were not indicative of issues at the Data Providers', but of issues with outbound ECHO network connectivity. These outages resulted in some delayed order dispatching, but had no other negative affects. On the morning of Tuesday, 3/16/10, Operations reported that they were having consistent issues accessing their tools box. There were no external impacts, but this raised the team's level of awareness that there were network issues. ECHO System Administrators were looking into this issue throughout the day.

When the network connectivity issues were identified on 3/16/10, it was initially assumed that the issues were related to the DDoS attack. However, further analysis identified that the issue was caused by a spanning tree protocol (STP) recalculation. ECHO System Administrators were able to identify a potential configuration issue between ECHO and EBNET relating to the redundant link which was installed early in the year. The issue was “self-correcting” but would periodically re-occur. During each recalculation, network connectivity would be dropped resulting in a loss of access to ECHO and WIST. Subsequent analysis has identified that the STP recalculations began on 1/27/10 and continued through 3/16/10. This analysis identifies that the network outages ECHO had seen over the past 2 months were related to STP recalculations, in addition to the DDoS. It is unclear why the spanning tree changes caused such a large impact to the Operational RAC nodes on 3/16/10.

Monitoring of all RAC services was not configured on the new Operational hardware at the time of this event. Testing was being performed on the workload RAC system and had not yet been completed. Had automated monitoring been in place, the issues with the RAC nodes would have been identified sooner, however the problems encountered would not have been avoided. Once the RAC network issues were identified, the nodes were individually restarted and the ECHO system brought back on line. There were no issues bringing up the system.

Timeline:

- 14:42 EST – Network disturbance occurred.
- 15:00 EST – ECHO System Monitoring began reporting Operational API & WIST unavailability
- 16:00 EST – ECHO SAs, on-site at GSFC, reported seeing a lack of network traffic through the McAfee Intrusion Detection System (IDS).
- 18:30 EST – Network connectivity was restored to the ECHO internal network.
- 19:00 EST – ECHO DBAs reported seeing network errors on a one of the RAC cluster nodes.
- 20:30 EST – ECHO SAs notified that the Operational system issues continued.
- 21:00 EST – Removed the McAfee IDS
- 21:00 EST – Identified RAC interconnectivity issues and started to bring down all RAC nodes.
- 21:30 EST – All Oracle RAC nodes shutdown, started reboot of each node.
- 22:00 EST – All Oracle RAC nodes started and working properly.
- 22:00 EST – Both ECHO Kernels started. (ECHO API & WIST fully available)
- 22:45 EST – Ingest started. (All system components restored)

Associated Tickets/NCRs:

- ECHO_SA_TTs –
    1. 14000490 – Initial connectivity issues to Operations tools box.
    2. 14000500 – TT tracking the issues that occurred on 3/16 and their final resolution.
- ECHO_TTs – None
- ECHO_NCRs – None

Future Mitigation:

The ECHO System Administration team is working closely with the GSFC network team to resolve the configuration issues discovered between network components. Configuration changes were made to the ECHO network uplink during the 3/24/10 preventative maintenance. These changes did not resolve the issue, but improved the stability of network components to reduce the likelihood of a repeat of the issues seen on 3/16/10. To mitigate the issues seen on 3/16/10 entirely, discussions have been held with the GSFC network team and the corrective configuration changes are scheduled to be made during the 3/31/10 preventative maintenance.

Additional monitoring of the new Operational RAC nodes has been added to ECHO's system monitoring tool. This monitoring will help identify issues with RAC services in a more timely fashion.